*Rank Methods for Combination of Independent Experiments in Multivariate Analysis of Variance. Part One. Two Treatment Multiresponse Case**

PRANAB KUMAR SEN, *The University of North Carolina at Chapel Hill and University of Calcutta*

1. INTRODUCTION AND SUMMARY

In statistical experiments involving a set of multiresponse treatments and replicated under appreciably varied conditions, the assumptions underlying the usual parametric multivariate analysis of variance (MANOVA) procedures often appear to be quite stringent and dubious. The object of the present investigation is to propose and study a class of nonparametric MANOVA procedures, which not only remain valid for a broad class of parent distributions but also leave scope for the variation of the above distribution from one replicate to another in any arbitrary manner. The performance characteristics of the proposed

* Supported in part from Research Grant, GM-12868, from the National Institutes of Health, Public Health Services.

methods are also compared with that of the standard parametric methods.

Suppose, we have N subjects (plots) available for the experiment which are divided into r replicates (blocks) of sizes $N_1, \ldots, N_r$ respectively, so that $N_1+\ldots+N_r = N$. The ith replicate containing N_i subjects are further subdivided into two subgroups of sizes N_{i1} and N_{i2} respectively, and these two subgroups are treated with two different treatments A and B. Thus,

$$(1.1) \qquad N_i = N_{i1}+N_{i2}, \quad i = 1, \ldots, r, \ r \geqslant 2.$$

The application of any treatment is followed by a quantitative p variate ($p \geqslant 2$) response, which is a stochastic variable $\mathbf{X}_{ik,\alpha}$ where

$$(1.2) \qquad \mathbf{X}_{ik,\alpha} = (X_{ik,\alpha}^{(1)}, \ldots, X_{ik,\alpha}^{(p)}), \ \alpha = 1, \ldots, N_{ik}, \quad k = 1, 2, \quad i = 1, \ldots, r.$$

It is assumed that $\mathbf{X}_{ik,\alpha}$, $\alpha = 1, \ldots, N_{ik}$ are N_{ik} independent and identically distributed random variables (i.i.d.r.v.) distributed according to a continuous p-variate cumulative distribution function (c.d.f.) $F_{ik}(\mathbf{x})$, where it is given that

$$(1.3) \qquad F_{i2}(\mathbf{x}) = F_i(\mathbf{x}+\boldsymbol{\theta}), \ F_{i1}(\mathbf{x}) = F_i(\mathbf{x}), \quad i = 1, \ldots, r;$$

$\boldsymbol{\theta}$ being a real p-vector. In the sequel, it will be assumed that $r, p \geqslant 2$.

In MANOVA, the standard parametric tools assume that F_i, in (1.3), is a p-variate normal c.d.f. for each $i = 1, \ldots, r$, and further all these r c.d.f.'s have a common dispersion matrix $\boldsymbol{\Sigma}$. In fact, these MANOVA procedures are appreciably sensitive to any departure from either of the two assumptions made above, and hence, may be regarded to have only a very limited scope of applicability. In many cases, the different batches of subjects in the different replicates may be quite heterogeneous, and may even be from appreciably different populations. Here, we shall be concerned with the following two problems, where we make no assumption regarding the specific forms of $F_1, \ldots, F_r$ in (1.3).

First, referred to (1.3) we desire to test the null hypothesis

$$(1.4) \qquad H_0 : \boldsymbol{\theta} = \mathbf{0},$$

against the set of alternatives that $\boldsymbol{\theta}$ is a non-null p-vector. Second, we may also desire to estimate $\boldsymbol{\theta}$ (by a point as well as region value), when we have the reasons to believe that $\boldsymbol{\theta}$ is non-null. For both purposes, we have proposed and studied appropriate nonparametric MANOVA procedures. The beauty of the proposed method is that it not only remains valid for any continuous p variate c.d.f., but also allows $F_1, \ldots, F_r$, in (1.3), to be arbitrarily different from each other. The findings of this paper generalize some of the single replicate nonparametric MANOVA tests of Chatterjee and Sen ([2], [3]), Sen ([13], [14]) and Puri and Sen [10] to the multireplicate case. The nonparametric estimation procedure considered here generalizes the technique of Hodges and Lehmann [9] and Sen [11] not only to the multivariate but also to the multireplicate case. This may also be regarded as a direct multivariate generalization of a similar univariate nonparametric procedure considered by the present author ([12], [16]).

2. REPLICATED NONPARAMETRIC MANOVA TESTS

We pool the N_i observations of the ith replicate into a pooled sample of size N_i. Then with respect to the jth variate values $X^{(j)}_{ik,\alpha}$, $\alpha = 1, \ldots, N_{ik}$, $k = 1, 2$, we arrange the N_i observations in order of magnitude and denote the rank of $X^{(j)}_{ik,\alpha}$, in this set, by $R^{(j)}_{ik,\alpha}$ for $\alpha = 1, \ldots, N_{ik}$, $k = 1, 2$. So that $R^{(j)}_{i1,1}, \ldots, R^{(j)}_{i2,N_{i2}}$ is a permutation of the N_i numbers $1, \ldots, N_i$. By virtue of the assumed continuity of F_i, in (1.3), the possibility of ties may be ignored, in probability. The above ranking is done separately for each $j = 1, \ldots, p$, within each of the r replicates. Thus, the observed variate values in (1.2) are mapped into r independent sets of p-tuplet rank values

$$(2.1) \qquad \mathbf{R}_{ik,\alpha} = (R^{(1)}_{ik,\alpha}, \ldots, R^{(p)}_{ik,\alpha}), \alpha = 1, \ldots, N_{ik}, \ k = 1, 2; \quad i = 1, \ldots, r,$$

Let

$$(2.2)\qquad \boldsymbol{R}_i = (\boldsymbol{R}_{i1,\,1}, \ldots, \boldsymbol{R}_{i_2 N_{i_2}}), \quad i = 1, \ldots, r;$$

$$(2.3)\qquad \boldsymbol{R} = (\boldsymbol{R}_1, \ldots, \boldsymbol{R}_r).$$

Then $\boldsymbol{R}$ is a $p \times N$ matrix which is partitioned into r submatrices of orders $p \times \mathrm{N}_i$, $i = 1, \ldots, r$. Each of these submatrices is stochastic in nature and contains random rank p-tuplets. Thus $\boldsymbol{R}$ is a collection of r independent sets of random rank p-tuplets and will be termed the *compound collection matrix,* and $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_r$ as the *component collection matrices.*

Now for each combination of $(i, j,\ i = 1, \ldots, r,\ j = 1, \ldots, p)$, and for every positive integer N, let us define a sequence of elements (which are known functions of N)

$$(2.4)\qquad \boldsymbol{E}_N^{(i,j)} = (E_{N1}^{(i,j)}, \ldots, E_{NN}^{(i,j)}), \quad j = 1, \ldots, p, \quad i = 1, \ldots, r;$$

the rp different sequences need not be identical. However, in the majority of cases, we would prefer having $\boldsymbol{E}_N^{(i,j)} = \boldsymbol{E}_N$ for all i, j. For the time being, we assume that

$$(2.5)\qquad \frac{1}{N} \sum_{\alpha=1}^{N} |E_{N\alpha}^{(i,j)}|^{2+\delta} < \infty, \text{ for some } \delta > 0, \text{ and all N},$$

and later, we shall impose some further conditions on $\boldsymbol{E}_N^{(i,j)}$. Then, let $\boldsymbol{Z}_{N_i\beta}^{(j)} = 1$, if the βth smallest observation on the jth variate values in the ith replicate is from the treatment A, and let $\boldsymbol{Z}_{N_i\beta}^{(j)} = 0$, otherwise; for $\beta = 1, \ldots, N_1, j = 1, \ldots, p, i = 1, \ldots, r$. We also denote by

$$(2.6)\qquad \boldsymbol{Z}_{N_i}^{(j)} = (Z_{N_i}^{(j)}, \ldots, Z_{N_i N_i}^{(j)}), \quad j = 1, \ldots, p, \quad i = 1, \ldots, r$$

and consider the rp random variables defined as

$$(2.7)\qquad T_{N_i}^{(j)} = (\boldsymbol{Z}_{N_i}^{(j)} . \boldsymbol{E}_{N_i}^{(i,j)}) / (\boldsymbol{Z}_{N_i}^{(j)} . \boldsymbol{Z}_{N_i}^{(j)})$$

$$= \frac{1}{N_{i1}} \sum_{\beta=1}^{N_i} Z_{N_i\beta}^{(j)} E_{N_i\beta}^{(j)}, \quad j = 1, \ldots, p, \quad i = 1, \ldots, r.$$

Further, let

$$(2.8)\qquad T_{N\cdot}^{(j)} = \sum_{i=1}^{r} [N_{i1}(N_i-1)/N_i]T_{N_i}^{(j)},\ j = 1, \ldots, p$$

$$(2.9)\qquad \boldsymbol{T}_{N\cdot} = (T_{N\cdot}^{(1)}, \ldots, T_{N\cdot}^{(p)}).$$

Our proposed test is then based on $\boldsymbol{T}_{N\cdot}$. Before we present the test statistics and the test procedure, we consider the rationality of the test.

When the null hypothesis (1.4) is true, $\boldsymbol{X}_{ik,\alpha}$ $\alpha = 1, \ldots, N_{ik}$, $k = 1, 2$ are i.i.d.r.v. distributed according to the c.d.f. $F_i(\boldsymbol{x})$, defined in (1.3). Thus, by an adoptation of the rank-permutation argument of Chatterjee and Sen ([2], [3]) we may conclude that given the rank collection matrix $\boldsymbol{R}_i$, in (2.2), the conditional distribution over the $N_i!$ permutations of the columns of $\boldsymbol{R}_i$ would be uniform under H_0 in (1.4), whatever the c.d.f. $F_i(\boldsymbol{x})$, may be. Consequently, given $\boldsymbol{R}_i$, all possible partitionings of the N_i rank p-tuplets into two subsets of sizes N_{i1} and N_{i2} respectively are equally likely (conditioned on the given $\boldsymbol{R}_i$) under H_0 in (1.4), and the permutational probability measure for each such partitioning is equal to $\binom{N_i}{N_{1i}}^{-1}$. We now consider the product permutational probability measure induced by the r independent sets of partitionings arising out of the r component collection matrices $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_r$. Evidently, this is given by

$$(2.10)\qquad \prod_{i=1}^{r} \binom{N_i}{N_{i1}}^{-1} = 1/N^* \text{ (say).}$$

Hence, conditioned on the component collection matrices of the compound collection matrix $\boldsymbol{R}$ to be all given, there are in all N^* possible partitionings and each such partitioning has (conditionally) a common permutational probability $1/N^*$, when H_0 in (1.4) is true. Thus, if we denote this permutational probability measure by $\mathcal{P}(\boldsymbol{R})$, and consider the permutation distribution of $\boldsymbol{T}_{N\cdot}$ (defined in (2.9),) induced by $\mathcal{P}(\boldsymbol{R})$, then any test function $\phi(\boldsymbol{T}_{N\cdot})$ based on this permutation distribution of $\boldsymbol{T}_{N\cdot}$ will have a strictly distribution-free structure when H_0 in (1.4) is true. The proposed test is thus a permutation test and is based on an extended rank-permutation argument.

To construct the test statistic S_N, we define first

$$\bar{E}_{N_i}^{(i,j)} = \frac{1}{N_i} \sum_{\beta=1}^{N_i} E_{N_i\beta}^{(i,j)}, \quad j = 1, \ldots, p, \quad i = 1, \ldots, r, \tag{2.11}$$

$$\bar{E}_N^{(j)} = \sum_{i=1}^{r} [N_{i1}(N_i-1)/N_i] \bar{E}_{N_i}^{(i,j)}, \quad j = 1, \ldots, p; \tag{2.12}$$

$$\bar{\boldsymbol{E}}_N = (\bar{E}_N^{(1)}, \ldots, \bar{E}_N^{(p)}). \tag{2.13}$$

Also, let

$$v_{jl}^{(i)} = \frac{1}{N_i-1} \left\{ \sum_{k=1}^{2} \sum_{\alpha=1}^{N_{ik}} E_{ik,\alpha}^{(j)} E_{ik,\alpha}^{(l)} - N_i \bar{E}_{N_i}^{(i,j)} \bar{E}_{N_i}^{(i,l)} \right\}, \tag{2.14}$$

for $j, l = 1, \ldots, p$, $i = 1, \ldots, r$, where $E_{ik,\alpha}^{(j)}$ is the value of $E_{N_i}^{(i,j)}$ associated with the value of $\beta = R_{ik,\alpha}^{(j)}$ for $\alpha = 1, \ldots, N_{ik}$, $k = 1, 2$, $j = 1, \ldots, p$; $i = 1, \ldots, r$. Further, let

$$n_i = N_{i1}N_{i2}/N_i, \quad i = 1, \ldots, r \text{ and } n = \sum_{i=1}^{r} n_i; \tag{2.15}$$

$$v_{jl}(\boldsymbol{R}) = \sum_{i=1}^{r} n_i v_{jl}^{(i)}/n, \ j, l = 1, \ldots, p. \tag{2.16}$$

Then, it can be shown following a few simple steps that

$$E\{\boldsymbol{T}_{N\cdot} \mid \mathscr{P}(\boldsymbol{R})\} = \bar{\boldsymbol{E}}_N, \tag{2.17}$$

$$E\{\boldsymbol{T}_{N\cdot} - \bar{\boldsymbol{E}}_N)'(\boldsymbol{T}_{N\cdot} - \bar{\boldsymbol{E}}_N) \mid \mathscr{P}(\boldsymbol{R})\} = n\boldsymbol{V}(\boldsymbol{R}),$$

where

$$\boldsymbol{V}(\boldsymbol{R}) = ((v_{jl}(\boldsymbol{R})))_{j,l=1,\ldots,p}. \tag{2.18}$$

we now let $\boldsymbol{V}^{-1}(\boldsymbol{R}) = ((v^{jl}(\boldsymbol{R})))_{j,l=1,\ldots,p}$.

Following then essentially the same argument as in Puri and Sen [10], we may consider the following test-statistic

$$S_N = \frac{1}{n}(\boldsymbol{T}_{N\cdot} - \bar{\boldsymbol{E}}_N)\boldsymbol{V}^{-1}(\boldsymbol{R})(\boldsymbol{T}_{N\cdot} - \bar{\boldsymbol{E}}_N)', \tag{2.19}$$

which is a positive semidefinite quadratic form in $(\boldsymbol{T}_{N\cdot}-\overline{\boldsymbol{E}}_N)$. S_N will be small only when $(\boldsymbol{T}_{N\cdot}-\boldsymbol{E}_N)$ consists of elements of small magnitude. When H_0 in (1.4) is true, S_N will have have N^* possible (permuted) values (not necessarily all distinct), and the permutational probability measure attached to each of these points is $1/N^*$. On the other hand, if H_0 is not true i.e., $\boldsymbol{\theta} \neq \boldsymbol{0}$ (referred to (1.3), then it can be shown that by proper choice of $\boldsymbol{E}_{N_i}^{(i,j)}$, $i=1, \ldots, r$, $j = 1, \ldots, p$, S_N can be made large, in probability. Thus, it seems reasonable to base our permutation test on S_N, using the right hand tail of the permutation c.d.f. of S_N as the appropriate critical region. Hence, we propose the following test function $\phi(S_N)$:

$$(2.20) \qquad \phi(S_N) = \begin{cases} 1, & \text{, if } S_N > S_{N,\epsilon}, \\ a_{N,\epsilon}, & \text{if } S_N = S_{N,\epsilon}, \\ 0, & \text{, if } S_N < S_{N,\epsilon}, \end{cases}$$

where $S_{N,\epsilon}$ and $a_{N,\epsilon}$ are so chosen that

$$(2.21) \qquad E\{\phi(S_N) \mid \mathscr{P}(\boldsymbol{R})\} = \epsilon : 0 < \epsilon < 1,$$

ϵ being the preassigned level of significance of the test. Note that (2.21) implies that $E\{\phi(S_N) \mid H_0\} = \epsilon$. So that $\phi(S_N)$ is a strictly size ϵ test. The values of $S_{N,\varepsilon}$ and $a_{N,\varepsilon}$ depend on the particular $\boldsymbol{R}$. In small samples, their values are to be evaluated from the exact permutation c.d.f. of S_N. The problem though appears to be deterministic, the labor involved in this numerical evaluation increases prohibitively with the increase in the sample sizes. In view of this, we present below the asymptotic permutation theory related to this problem and later, we shall see how the same can be used to simplify the large sample approach to this permutation test procedure.

Extending the ideas of Chernoff and Savage [4] (see also [5], [10]) to the multireplicate multivariate case, we write

$$(2.22) \qquad E_{N_i\alpha}^{(i,j)} = J_{N_i}^{(i,j)}(\alpha/(N_i+1)), \; 1 \leqslant \alpha \leqslant N_i, \quad j = 1, \ldots, p$$
$$i = 1, \ldots, r,$$

where $J_{N_i}^{(i,j)}$ need be defined only at $\alpha/(N_i+1)$, $\alpha = 1, \ldots, N_i$ and its domain of definition may be extended to the open interval (0, 1) in any conventional manner (cf. [10]). Let then

$$(2.23)\quad F_{N_i,k}^{(j)}(x) = \frac{1}{N_{ik}}\ \{\text{number of } X_{ik,\alpha}^{(j)} \leqslant x\},\quad j = 1, \ldots, p,\quad i = 1, \ldots, r;\ k = 1, 2;$$

$$(2.24)\quad H_{N_i}^{(j)}(x) = \frac{1}{N_i}\{N_{i1}F_{N_i,1}^{(j)}(x)+N_{i2}F_{N_i,2}^{(j)}(x)\},\quad j = 1, \ldots, p,\quad i = 1, \ldots, r.$$

The marginal c.d.f. of $X_{ik,\alpha}^{(j)}$ is denoted by $F_{i,k}^{(j)}(x)$ for $k = 1, 2$, and let

$$(2.25)\qquad H_i^{(j)}(x) = \frac{1}{N_i}\{N_{i1}F_{i,1}^{(j)}(x)+N_{i2}F_{1,2}^{(j)}(x)\},\quad i = 1, \ldots, r,\quad j = 1, \ldots, p.$$

Similarly, let

$$(2.26)\quad F_{N_i,k}^{(j,l)}(x, y) = \frac{1}{N_{ik}}\{\text{number of } (X_{ik,\alpha}^{(j)}, X_{ik,\alpha}^{(l)}) \leqslant (x, y)\},\quad k = 1, 2,;$$

$$(2.27)\qquad H_{N_i}^{(j,l)}(x, y) = \frac{1}{N_i}\{N_{i1}F_{N_i,1}^{(i,l)}(x, y)+N_{i2}F_{N_i,2}^{(i,l)}(x, y),\},\quad j \neq l = 1, \ldots, p,\quad i = 1, \ldots, r.$$

Also, the joint (marginal) c.d.f. of $X_{ik,\alpha}^{(j)}, X_{ik,\alpha}^{(l)}$ is denoted by $F_{i,k}^{(j,l)}(x, y)$ for $k = 1, 2$, and we let

$$(2.28)\qquad H_i^{(j,l)}(x, y) = \frac{1}{N_i}\{N_{i1}F_{i,1}^{(j,l)}(x, y)+N_{i2}F_{1,2}^{(j,l)}(x, y)\},\quad j \neq l = 1, \ldots, p,\quad i = 1, \ldots, r.$$

It may be noted that if we define

$$(2.29)\qquad \lambda_{N_i}^{(i)} = N_{i1}/N_i,\qquad i = 1, \ldots, r,$$

then $H_i^{(j)}$ and $H_i^{(j,l)}$ both depend explicitly on $\lambda_{N_i}^{(i)}$, $i = 1, \ldots, r$. We now impose the following conditions:

(C.1) $\lambda_{N_1}^{(1)}, \ldots, \lambda_{N_r}^{(r)}$ are regarded as fixed and are all bounded away from zero and one.

(C.2) $J^{(i,j)}(H) = \lim_{N_i \to \infty} J_{N_i}^{(i,j)}(H)$ exists for all $0 < H < 1$ and is not a constant. For our purpose, we impose further that $J^{(i,j)}(H)$ is monotonic in H, for all $i = 1, \ldots, r$, $j = 1, \ldots, p$.

(C.3) If $I_{N_i}^{(j)} = \{x : 0 < H_{N_i}^{(j)}(x) < 1\}$, then

$$\int_{I_{N_i}^{(j)}} \left[J_{N_i}^{(i,j)}\left(\frac{N_i}{N_i+1} H_{N_i}^{(j)}\right) - J^{(i,j)}\left(\frac{N}{N_i+1} H_{N_i}^{(j)}\right)\right] dF_{N_j,1}^{(j)}(x)$$

$$= o_p(N_i^{-\frac{1}{2}}), \text{ for all } j = 1, \ldots, p \text{ and } i = 1, \ldots, r.$$

(C.4) $$\left| \frac{d^l}{dH^l} J^{(i,j)}(H) \right| \leqslant K[H(1-H)]^{-l-\frac{1}{2}+\delta},$$

for $l = 0, 1$ and some $\delta > 0$.

(C.5) $$\int_{I_{N_i}^{(j)} \times I_{N_i}^{(l)}} \int \left(J_{N_i}^{(i,j)}\left[\frac{N_i}{N_i+1} H_{N_i}^{(j)}\right) J_{N_i}^{(i,l)}\left(\frac{N_i}{N_i+1} H_{N_i}^{(l)}\right) -$$

$$J^{(i,j)}\left(\frac{N_i}{N_i+1} H_{N_i}^{(j)}\right) J^{(i,l)}\left(\frac{N_i}{N_i+1} H_{N_i}^{(l)}\right)\right] dH_{N_i}^{(j,l)}(x, y) = o_p(1),$$

for all $j, l = 1, \ldots, p$, $i = 1, \ldots, r$.

(C.6) Let us define

(2.30) $$\nu_{jl}^{(i)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [J^{(i,j)}(H_i^{(j)}) J^{(i,l)}(H_i^{(l)})] dH_i^{(j,l)}(x, y)$$

$$-\left[\int_{-\infty}^{\infty} J^{(i,j)}(H_i^{(j)}) dH_i^{(j)}(x)\right]\left[\int_{-\infty}^{\infty} J^{(i,l)}(H_i^{(l)}) dH_i^{(l)}(y)\right],$$

for $j, l = 1, \ldots, p$, and

(2.31) $$\boldsymbol{\nu}^{(i)} = ((\nu_{jl}^{(i)}))_{j,l=1,\ldots,p} \quad i = 1, \ldots, r.$$

Then $\boldsymbol{\nu}^{(i)}$ is positive definite for all $i = 1, \ldots, r$. We shall see later on that conditions (C.1) through (C.6) are less restrictive than the ones required in the parametric case. Conditions (C.1) through (C.4) are the same as in the Chernoff-Savage theorem ([4], also [5]), while conditions (C.5) is required to study the convergence of the matrix $\boldsymbol{V}(\boldsymbol{R})$ (defined in (2.18)), and the condition (C.6) is required to assume that $\boldsymbol{V}(\boldsymbol{R})$ has a positive definite limit, in probability. Let us define now

$$\boldsymbol{\nu} = \sum_{i=1}^{r} n\, \boldsymbol{\nu}^{(i)}/n \tag{2.32}$$

where $n_1, \ldots, n_r$ and n are defined in (2.15). It then follows from (C.6), (2.31) and (2.32) that $\boldsymbol{\nu}$ is also positive definite.

Theorem 2.1. *For arbitrary continuous c.d.f.'s* $F_1, \ldots, F_r$ *and any real and finite* $\boldsymbol{\theta}$ *(in* (1.3),*) under the conditions* (C.1) *through* (C.5), $\boldsymbol{V}(\boldsymbol{R}) \xrightarrow{P} \boldsymbol{\nu}$.

Outline of proof. If follows from Theorem 4.2 of Puri and Sen [10] (which essentially relates to the unireplicate case) that under the stated regularity conditions

$$v_{jl}^{(i)} \xrightarrow{P} \nu_{jl}^{(i)} \text{ for all } j, l = 1, \ldots, p; \quad i = 1, \ldots, r. \tag{2.33}$$

Using then (2.16), (2.30) and (2.32), we get after some simple algebraic manipulations that

$$v_{jl}^{(i)} \xrightarrow{P} \nu_{jl} \text{ for all } j, l = 1, \ldots, p.$$

Hence, the theorem.

Theorem 2.2. *Under the conditions* (C.1), $i = 1, 2, 3, 4$ *and* $\boldsymbol{V}(\boldsymbol{R})$ *being positive definite, the permutation distribution of* S_N *(defined in* (2.19),*) asymptotically, reduces to a chi-square distribution with* p *degrees of freedom.*

Outline of proof. It follows from Theorem 5.1 of Puri and Sen [10] that under the stated regularity conditions, the joint permutation distribution of $N_i^{\frac{1}{2}}[T_{N_i}^{(j)} - \bar{E}_{N_i}^{(i,j)}]$, $j = 1, \ldots, p$ asymptotically reduces to a p-variate normal distribution. Now, the sets $N_i^{\frac{1}{2}}[T_{N_i}^{(j)} - \bar{E}_{N_i}^{(i,j)}]$, $j = 1, \ldots, p$ $(i = 1, \ldots, r)$ being all stochastically independent, it follows from (2.8) that the joint permutation distribution of $N^{-\frac{1}{2}}[\boldsymbol{T}_{N\cdot}-\boldsymbol{E}_N]$ can be obtained by convolution of r independent and asymptotically multinormal distributions. Hence, it can be easily shown that $n^{-\frac{1}{2}}\{\boldsymbol{T}_{N\cdot}-\boldsymbol{E}_N\}$ has a permutation distribution, which is asymptotically multinormal if $\boldsymbol{V}(\boldsymbol{R})$ is positive definite. Again, if $\boldsymbol{V}(\boldsymbol{R})$ is positive definite, the above permutation distribution is essentially non-singular and hence, applying a well-known result on the asymptotic distribution of quadratic forms associated with asymptotic normal distributions (cf. Sverdrup [17]), we arrive at the desired result that S_N has asymptotically a chi-square (permutation) distribution with degrees of freedom.

Hence, the theorem.

By virtue of Theorem 2.1 and condition (C.6), we get that $\boldsymbol{V}(\boldsymbol{R})\xrightarrow{P}\boldsymbol{\nu}$, which is positive definite. Hence, from the preceeding theorem, we arrive at the following.

Theorem 2.3. *Under conditions* (C.1), $i=1, \ldots, 6$, *the permutation distribution of* S_N *reduces asymptotically in probability, to a* χ^2 *distribution with p.d.f.*

Thus, it follows from Theorem 2.3 that the test $\phi(S_N)$ in (2.20) asymptotically, in probability, reduces to

$$\phi(S_N) = 1, \text{ if } S_N \geqslant \chi^2_{p,\varepsilon} \qquad (2.34)$$
$$= 0, \text{ if } S_N < \chi^2_{p,\varepsilon},$$

where $\chi^2_{p,\epsilon}$, is the $100(1-\epsilon)\%$ point of a χ^2 distribution with p.d.f. Using condition (C.2), it can be shown with little difficulty that $\frac{1}{N}\{\boldsymbol{T}_{N\cdot}-\overline{\boldsymbol{E}}_N\}$ converges to a non-null vector if $\boldsymbol{\theta}$ (in (1.3),) is non-null. Consequently, using (2.19) and Theorem 2.3, it is easy to show that the test $\phi(S_N)$ in (2.20) or (2.34) is consistent against any $\boldsymbol{\theta}\neq\boldsymbol{0}$.

Now, as in most of the non-parametric tests in MANOVA, the study of the exact power of $\phi(S_N)$ seems to be considerably difficult and the same depends heavily on the parent c.d.f.'s $F_1, \ldots, F_r$ and the particular $\boldsymbol{\theta}$. No simple expression can be attached to such an exact power-function. However, we are in a position to study the asymptotic power properties of the test $\phi(S_N)$, and this we do for a sequence of alternatives $\{\boldsymbol{\theta}_N\}$, so chosen that $E\{\phi(S_N)|\theta_N\}$ tends to a finite limit $\gamma : \epsilon < \gamma < 1$ as $N \to \infty$. We now define n as in (2.15), and note that n is an increasing function of N_{ik}, $k = 1, 2$, $i = 1, \ldots, r$, and it tends to ∞ as $N \to \infty$, subject to (C.1). Then, we specify the sequence of alternative hypotheses $\{H_N\}$ by

$$H_N : \boldsymbol{\theta}_N = n^{-\frac{1}{2}} . \boldsymbol{\delta}, \tag{2.35}$$

where $\boldsymbol{\delta}$ is any real and finite p-vector. We also define

$$a_{ij} = \int_{-\infty}^{\infty} \frac{d}{dx} J^{(i,j)}(H_i^{(j)}(x))dH_i^{(j)}(x), \quad i = 1, \ldots, r, \ j = 1, \ldots, p \tag{2.36}$$

$$\eta_{j,N} = \left(\sum_{i=1}^{r} n_i a_{ij}/n \right) \delta_j, \ j = 1, \ldots, p;$$

$$\eta_N = (\eta_{1,N}, \ldots, \eta_{p,N}) \tag{2.37}$$

At this stage, we require to put some restrictions on $n_1, \ldots, n_r$, in (2.15). We assume that as $n \to \infty$, $n_i/n \to p_i$, $i = 1, \ldots, r$ where $p_1, \ldots, p_r$ are all bounded away from zero and one and they add up to unity. Under this condition η_N, in (2.37), tends to a limit-vector, which is denoted by $\boldsymbol{\eta}$.

Theorem 2.4. *Under conditions* (C.1), $i = 1, \ldots, 6$ *and the one on n_i's stated above S_N has asymptotically, under $\{H_N\}$, a non central χ^2 distribution with p.d.f. and the noncentrality parameter is*

$$\Delta_S = \boldsymbol{\eta} \boldsymbol{\nu}^{-1} \boldsymbol{\eta}'.$$

Proof: For arbitrary $F_1, \ldots, F_r$ and $\boldsymbol{\theta}$ in (1.3) (not necessarily the sequence $\{\theta_N\}$), the joint asymptotic normality of $[N_i^{\frac{1}{2}}\{T_{N_i}^{(j)} - \bar{E}_{N_i}^{(i,j)}\}, \ j = 1, \ldots, p]$ follows readily (under conditions (C.1), $i = 1, \ldots, 4$) from Theorem 5.1 of Puri and Sen [10]. Thus, again by simple convolution, we get that for arbitrary $F_1, \ldots, F_r$ and $\boldsymbol{\theta}$, $N^{-\frac{1}{2}}[\boldsymbol{T}_{N\cdot}-\overline{\boldsymbol{E}}_N]$ has asymptotically a multinormal distribution, under conditions (C.i), $i = 1, \ldots, 4$. Now, under $\{H_N\}$ in (2.35), it follows by more or less routine computation that

$$E\{n^{-\frac{1}{2}}[\boldsymbol{T}_{N\cdot}-\boldsymbol{E}_N]/H_N\} \rightarrow \boldsymbol{\eta} \text{ as } N \rightarrow \infty, \tag{2.38}$$

and further it is also easily shown that

$$E\{n^{-1}[\boldsymbol{T}_{N\cdot}-\boldsymbol{E}_N]'[\boldsymbol{T}_{N\cdot}-\overline{\boldsymbol{E}}_N]/H_N\} \rightarrow \boldsymbol{\nu} \text{ as } N \rightarrow \infty, \tag{2.39}$$

where $\boldsymbol{\nu}$ is defined in (2.32). (Note that, in this case, in (2.30) and (2.36), $H_i^{(j)}$, $H_i^{(j,l)}$ are to be replaced by $F_i^{(j)}$ and $F_i^{(j,l)}$ respectively, for all $j, l = 1, \ldots, p$ and $i = 1, \ldots, r$). Consequently, by a well-known limit theorem, we get that under $\{H_N\}$

$$S_N^* = \frac{1}{n}\,[(\boldsymbol{T}_{N\cdot}-\overline{\boldsymbol{E}}_N)\nu^{-1}(\boldsymbol{T}_{N\cdot}-\boldsymbol{E}_N)'] \tag{2.40}$$

has asymptotically a noncentral χ^2 distribution with p.d.f. and the noncentrality parameter Δ_S, defined in the theorem. Finally, using Theorem 2.1 it is easy to show that under $\{H_N\}$ in (2.35), $\boldsymbol{V}(\boldsymbol{R}) \xrightarrow{P} \boldsymbol{\nu}$, where $\boldsymbol{V}(\boldsymbol{R})$ is defined in (2.18). Consequently, from (2.19), (2.35) and 2.40) we get after a few simple steps that under $\{\boldsymbol{H}_N\}$, $S_N \overset{P}{\approx} S_N^*$.

Hence, the theorem.

Let us now consider the standard parametric MANOVA test which is based upon the assumption that the c.d.f.'s $F_1, \ldots, F_r$ (in (1.3),) are all p-variate normal with a common covariance matrix $\boldsymbol{\Sigma}$. The test-statistic (say R_N) is essentially a likelihood ratio test criterion (c.f. Wilks [18, p 561]) and it is easily seen that under H_0, in (1.4), R_N has asymptotically a χ^2 distribution with p d.f., while under $\{H_N\}$, in (2.35), R_N has asymptotically a noncentral χ^2 distribution with p d.f. and the noncentrality parameter

$$\Delta_R = \boldsymbol{\delta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}'. \tag{2.41}$$

Thus the asymptotic efficiency (in the Pitman-sense) of the permutation MANOVA test based on S_N with respect to the standard parametric MANOVA test based on R_N comes out as

$$(2.42)\qquad e^{\{H_N\}}_{\{S_N\},\{R_N\}} = \boldsymbol{\eta}\boldsymbol{\nu}^{-1}\boldsymbol{\eta}'/\boldsymbol{\delta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}' = e(\boldsymbol{\delta}, \boldsymbol{J}, \boldsymbol{F}),$$

where $\boldsymbol{F} = (F_1, \ldots, F_r)$ and $\boldsymbol{J} = (J^{(1,1)}, \ldots, J^{(1,p)}, \ldots, J^{(r,1)}, \ldots, J^{(r,p)})$. Further, in the parametric case, if $F_1, \ldots, F_r$ are normal and have covariance matrices $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_r$ respectively, which are not all identical, the statistic R_N has no simple distribution for small samples. However, for large samples, it can be shown that under $\{H_N\}$, in (2.35), R has asymptotically a noncentral χ^2 distribution with p d.f. and the noncentrality parameter

$$(2.43)\qquad \Delta^*_R = \boldsymbol{\delta}\overline{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\delta}', \ \boldsymbol{\Gamma} = \sum_{i=1}^{r} n_i\Sigma_i/n;$$

(consequently, under H_0, R_N has asymptotically a χ^2 distribution with p d.f.). In this case, the asymptotic efficiency in (2.42) will have to be adjusted only by replacing $\boldsymbol{\Sigma}$ by $\overline{\boldsymbol{\Sigma}}$.

It may be noted in this connection that a second type of non-parametric MANOVA tests may also be constructed for the above purpose. This type of tests are somewhat restrictive in the sense that it allows $F_1, \ldots, F_r$ to be continuous p-variate c.d.f's but assumes that $F_1, \ldots, F_r$ have the same functional form apart from possible variation of the location vectors only. In a sense, it is thus essentially the parametric MANOVA test with the assumption of normality replaced by a broad family of continuous c.d.f's, but the assumption of identity of the covariance matrices being implicit. This type of tests is based upon rankings after alignment and may be regarded as a direct multivariate generalization of a similar class of univariate rank-tests considered by Hodges and Lehmann [8]. The study of such tests in the multivariate case poses some further problems, connected with rank-order tests, and because of the somewhat different type of approach, the solutions in this case and their lengthy deductions will be considered in a separate issue.

3. ESTIMATION OF $\boldsymbol{\theta}$ USING RANK-ORDER TESTS.

We shall now consider the problem of estimating $\boldsymbol{\theta}$ in (1.3), without assuming the functional forms of $F_1, \ldots, F_r$ or their co-variance-matrices to be all identical. For this, we write

(3.1) $\boldsymbol{X}_{ik}^{(j)} = \left(X_{ik,1}^{(j)}, \ldots, X_{ik,N_{ik}}^{(j)}\right), k=1, 2, j = 1, \ldots, p, i = 1, \ldots r;$

(3.2) $\boldsymbol{X}_{ik} = (\boldsymbol{X}_{ik}^{(1)}, \ldots, \boldsymbol{X}_{ik}^{(p)}),\ k = 1, 2,\ i = 1, \ldots, r;$

(3.3) $\boldsymbol{X}_{.k}^{(j)} = (\boldsymbol{X}_{1k}^{(j)}, \ldots, \boldsymbol{X}_{rk}^{(j)}),\quad k = 1, 2,\quad j = 1, \ldots, p;$

(3.4) $\boldsymbol{X}_{.k} = (\boldsymbol{X}_{.k}^{(1)}, \ldots, \boldsymbol{X}_{.k}^{(p)}),\quad k = 1, 2.$

Let $\boldsymbol{I}_m$ be an m-vector with unit elements, and we denote

(3.5) $(x_1+a, \ldots, x_m+a) = \boldsymbol{X}_m+a\boldsymbol{I}_m$

We then consider a rank-statistic $h(\boldsymbol{X}_m, \boldsymbol{Y}_n)$ which statisfies the following two conditions:

(a) $h(\boldsymbol{X}_m+a\boldsymbol{I}_m, \boldsymbol{Y}_n)$ is increasing in a,

(b) If the elements of $\boldsymbol{X}_m$ and $\boldsymbol{Y}_n$ are i.i.d.r.v., then $h(\boldsymbol{X}_m, \boldsymbol{Y}_n)$ has a strictly distribution-free structure, and the c.d.f of $h(\boldsymbol{X}_m, \boldsymbol{Y}_n)$ in this case is denoted by $G_0(h/m, n)$.

Let $\mu_{m,n}$ be any convenient measure of location of G_0, and as G_0 is distribution-free $\mu_{m,n}$ is a known function of (m, n). For the definition of $\mu_{m,n}$, we may adopt the convention in ([9], [16]). Let us then define

$$h_j^*(\boldsymbol{X}_{.1}^{(j)}+a_j\boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}^{(j)}) = \frac{1}{n} \sum_{i=1}^{r} n_i h_j(\boldsymbol{X}_{i1}^{(j)}+a_j\boldsymbol{I}_{N_{i1}}, \boldsymbol{X}_{i2}^{(j)}), \tag{3.6}$$

for $j = 1, \ldots, p$, where $n_1, \ldots, n_r$ and n are defined in (2.15). The appropriate location of the c.d.f. of $h_j^*(\boldsymbol{X}_{.1}^{(j)}+\theta_j\boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}^{(j)})$ can be easily derived from $\mu_{N_{i1}, N_{i2}}^{(j)}, N_2,\ i = 1, \ldots, r$ and is denoted by $\mu_N^{(j)},\ j = 1, \ldots, p$. So that $\boldsymbol{\mu}_N = (\mu_N^{(1)}, \ldots, \mu_N^{(p)})$ is a known p-vector. Let us again define

$$\hat{\theta}_{j.1} = \text{Inf}\,\{\theta : h_j^*(\boldsymbol{X}_{.1}^{(j)}+\theta\boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}^{(j)}) > \mu_N^{(j)}\}, \tag{3.7}$$

$$\hat{\theta}_{j.2} = \text{Sup}\,\{\theta : h_j^*(\boldsymbol{X}_{.1}^{(j)}+\theta\boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}^{(j)}) < \mu_N^{(j)}\};\ j = 1, \ldots, p. \tag{3.8}$$

Then our proposed estimate of $\boldsymbol{\theta}$ in (1.3) is

$$(3.9)\quad \hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_p);\ \hat{\theta}_j = (\hat{\theta}_{j.1}+\hat{\theta}_{j.2})/2, \quad j = 1, \ldots, p.$$

The proposed method is a generalization of a similar univariate method considered by Hodges and Lehmann [9] and Sen [11] not only to the multivariate but also to the multireplicate case. Incidently, this also generalizes Bickel's [1] results not only to the replicated two sample case but also to a more general class of test statistics. Finally, in the univariate replicated case, Sen ([12], [16]) has considered a similar method and the present one is a direct generalization of the same to the multivariate case.

Following then essentially the same technique as in [9], [16]) it is easily shown that if $F_1, \ldots, F_r$ are all continuous (absolutely continuous) so also is the estimate $\hat{\boldsymbol{\theta}}$. Further $\hat{\boldsymbol{\theta}}$ is translation invariant, i.e.,

$$(3.10)\qquad \hat{\boldsymbol{\theta}}(\boldsymbol{X}_{.1}+\boldsymbol{a}\boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}) = \hat{\boldsymbol{\theta}}(\boldsymbol{X}_{.1}, \boldsymbol{X}_{.2})+\boldsymbol{a},$$

for any real p-vector $\boldsymbol{a}$. Finally, if either $N_{i1} = N_{i2}$ for all $i = 1, \ldots, r$ or $F_1, \ldots, F_r$ are symmetric and if $\{h_j(\boldsymbol{X}_{i1}^{(j)}, \boldsymbol{X}_{i2}^{(j)})+h_j(\boldsymbol{X}_{i2}^{(j)}, \boldsymbol{X}_{i1}^{(j)}\}$ = constant (which may depend on j), for all $j = 1, \ldots, p$, then the distribution of $(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ is symmetric about $\mathbf{0}$.

We shall now consider the asymptotic properties of the estimate $\boldsymbol{\theta}$. We then define $n_1, \ldots, n_r$ and n as in (2.15), and let

$$(3.11)\qquad n_i/n = p_i : 0 < p_i < 1 \text{ for all } i = 1, \ldots, r.$$

Further, we impose an asymptotic condition *on* $h(\ )$; namely, we assume that under (1.3), the joint distribution of

$$(3.12)\quad [n_i^{\frac{1}{2}}\{h_j(\boldsymbol{X}_{i1}^{(j)}+(\theta_j+n_i^{-\frac{1}{2}}a_j)\boldsymbol{I}_{N_{i1}}, \boldsymbol{X}_{i2}^{(j)})-\mu_{N_{i1},N_{i2}}^{(j)}\},\ j = 1, \ldots, p]$$

is asymptotically p-variate normal with a mean vector $(a_1B_1^{(i)}(F_i), \ldots, a_pB_p^{(i)}(F_i))$ and a dispersion matrix $\boldsymbol{\nu}^{(i)}$, $i = 1, \ldots, r$. Let us then define

$$(3.13)\quad \bar{B}_j(F) = \sum_{i=1}^{r} p_iB_j^{(i)}(F_i),\ j = 1, \ldots, p,\ F = (F_1, \ldots, F_r);$$

$$(3.14) \qquad \boldsymbol{\nu} = \sum_{i=1}^{r} p_i \boldsymbol{\nu}^{(i)}, \ \boldsymbol{\tau} = ((\tau_{jl})),$$

$$\tau_{jl} = \nu_{jl} / \bar{B}_j(E).\bar{B}_l(\boldsymbol{F}), \ l, j = 1, \ldots, p.$$

Then essentially by an adaptation of the same technique as in ([9], [16]) with more or less straight forward generalization to the multivariate case, we arrive at the following theorem, where $\hat{\theta}$ in (3.9) is replaced by a sequence $\{\hat{\boldsymbol{\theta}}_n\}$ for the sequence of n.

Theorem 3.1. *Under the conditions stated above,* $n^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$ *has asymptotically a multinormal distribution with a null mean vector and a dispersion matrix* $\boldsymbol{\tau}$, *defined in* (3.14).

It may be noted here that if we now work with the class of rank-order tests $T_{N_i}^{(j)}$, $i = 1, \ldots, r$, $j = 1, \ldots, p$ and assume the conditions (C.i), $i = 1, 2, 3, 4$ of section 2 to hold, then the conditions (a) and (b) (of this section) imposed on $h(\ )$ are also satisfied. Further, in this case, the quantity $B_j^{(1)}(F_i)$ reduces to a_{ij}, defined in (2.36). Thus, the same class of rank-order tests may also be used in estimation problem.

In the parametric case, the conventional estimate of θ_j is $t_j = \sum_{i=1}^{r} p_i \bar{Z}_i^{(j)}$, $j = 1, \ldots, p$, where $\bar{Z}_i^{(j)}$ is the observed difference of means of $X_{i1,\alpha}^{(j)}$ and $X_{i2,\alpha}^{(j)}$, $j = 1, \ldots, p$, $i = 1, \ldots, r$. If $F_1, \ldots, F_r$ have a common covariance matrix Σ, it is easily seen that the covariance matrix of $n^{\frac{1}{2}}(\boldsymbol{t} - \boldsymbol{\theta})$ is also $\boldsymbol{\Sigma}$, (where $\boldsymbol{t} = (t_1, \ldots, t_p)$). If $F_1, \ldots, F_r$ have covariance matrices $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_r$, then if we define $\bar{\boldsymbol{\Sigma}}$ as in (2.43), the covariance matrix of $n^{\frac{1}{2}}(\boldsymbol{t} - \boldsymbol{\theta})$ is $\bar{\boldsymbol{\Sigma}}$. The asymptotic multi-normality of this vector estimate follows readily from the Central Limit Theorems. Thus, if we employ the generalized variance as a measure of efficiency, the asymptotic efficiency of the estimate (3.9) with respect to the parametric estimate $\boldsymbol{t}$ comes out as

$$(3.15) \qquad |\bar{\boldsymbol{\Sigma}} \,.\, \boldsymbol{\tau}^{-1}|^{1/p} = \{|\bar{\boldsymbol{\Sigma}}| / |\boldsymbol{\tau}|\}^{1/p} = e(\boldsymbol{\tau}, \bar{\boldsymbol{\Sigma}}), \ \text{(say)}.$$

We can easily imagine the closeness of (2.42) and (3.15) and later we shall study these further.

Before that we consider the following problem of confidence region of $\boldsymbol{\theta}$ based on rank-order tests. In the univariate case, it has been shown by the present author (cf. [11], [16]) that distribution-free confidence interval for θ_j can be obtained from the distribution of $h_j^*(\boldsymbol{X}_{.1}^{(j)}+\theta_j\boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}^{(j)})$, for any $j = 1, \ldots, p$. However, there are certain difficulties associated with the simultaneous confidence region for $\boldsymbol{\theta}$. This is due to the fact that marginally each $h_j^*(\boldsymbol{X}_{.1}^{(j)}+\theta_j \boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}^{(j)})$ has a nonparametric distribution for $j = 1, \ldots, p$. But, jointly these p statistics has a distribution-free structure only under permutation model, and the permutation covariance-matrix in this case depends on the unknown $\boldsymbol{\theta}$ and the given $\boldsymbol{X}_{.1}$, $\boldsymbol{X}_{.2}$. Thus, the permutation distribution depends on the unknown $\boldsymbol{\theta}$ in a somewhat involved manner, and it seems to be fairly difficult to suggest any procedure for very small samples. The procedure sketched below remains valid for moderately large samples, and may therefore be regarded as an asymptotic procedure.

For the confidence region problem, we assume that the rank-order tests of Section 2 are used and borrow therefore the same notations. We can estimate $\nu_{jl}^{(i)}$ from the ith replicate in the following manner. Since, under (1.3), F_{i1} and F_{i2} have the same functional form and they differ only by a location vector, if we adopt the ranking procedure in Section 2 for each of the two samples (separately) within the ith replicate, and for each sample we use an estimate essentially similar to the one in (2.14), then by combining these two estimates within the ith replicate (with weights equal to $N_{i1}-1$ and $N_{i2}-1$ respectively), we get an estimate of $\nu_{jl}^{(i)}$, for $j,\ l = 1, \ldots, p,\ i = 1, \ldots, r$. It can be readily shown using Theorem 5.2 of Puri and Sen [10] that these are all consistent estimates. Once these are obtained, we estimate ν_{jl} in (2.32), by the same linear function (i.e. (2.32)) in the estimates. We denote these $\hat{\nu}_{jl}$, $j, l = 1, \ldots, p$, and the matrix by $\hat{\boldsymbol{\nu}}$. Consequently, we replace (2.19) by

$$H_N(\boldsymbol{\theta}) = \frac{1}{n} \{(\boldsymbol{h}^*(\boldsymbol{\theta})-\boldsymbol{\mu}_N)\hat{\boldsymbol{\nu}}^{-1}(\boldsymbol{h}^*(\boldsymbol{\theta})-\boldsymbol{\mu}_N)'\}, \tag{3.16}$$

where

(3.17) $\boldsymbol{h}^*(\boldsymbol{\theta}) = (h_1^*(\boldsymbol{X}_{.1}^{(1)}+\theta_1 \boldsymbol{I}_{N.1}, \boldsymbol{X}_{.2}^{(1)}), \ldots, h_p^*(\boldsymbol{X}_{.1}^{(p)}+\theta\, I_{p.1}, \boldsymbol{X}_{.2}^{(p)}))$

Then, proceeding precisely on the same line as in Theorem 2.2, it can be shown that under (1.3) and asymptotically

(3.18) $$P\{H_N(\boldsymbol{\theta}) \leqslant \chi^2_{p,\varepsilon} \,|\, \boldsymbol{\theta}\} = 1-\epsilon,\ 0 < \epsilon < 10$$

where $\chi^2_{p,\varepsilon}$ is defined in (2.34). We now select $1-\epsilon$ as the desired confidence coefficient. Now $H_N(\boldsymbol{\theta}) \leqslant \chi^2_{p,\varepsilon}$ describes an ellipsoid in $\boldsymbol{h}^*(\boldsymbol{\theta})$ with origin $\boldsymbol{\mu}_N$. For any point $\boldsymbol{a}$ on the boundary of this ellipsoid, we have

(3.19) $$\boldsymbol{h}^*(\boldsymbol{\theta}) = \boldsymbol{a}.$$

We then solve the set of p equations in p unknowns in (3.19), and denote this solution as $\hat{\boldsymbol{\theta}}(\boldsymbol{a})$ where we adopt the following convention to achieve uniqueness of $\hat{\boldsymbol{\theta}}(\boldsymbol{a})$ for any $\boldsymbol{a}$.

(3.20) $$\hat{\boldsymbol{\theta}}(\boldsymbol{a}) = \{\boldsymbol{h}^*(\hat{\boldsymbol{\theta}}) = \boldsymbol{a} : \|\boldsymbol{\theta}(\boldsymbol{a})-\hat{\boldsymbol{\theta}}\| \text{ is maximum}\},$$

where $\hat{\boldsymbol{\theta}}$ is defined in (3.9), i.e., for any $\boldsymbol{a}$, we take the extreme (distant) value of $\boldsymbol{\theta}$ for which (3.19) holds. Now using the method of (3.19) and (3.20) and allowing $\boldsymbol{a}$ to assume all possible values on the boundary of the ellipsoid $H_N(\boldsymbol{\theta}) \leqslant \chi_{p,\varepsilon}$, it is easily seen (using the monotonicity of $\boldsymbol{h}^*(\boldsymbol{\theta})$ with $\boldsymbol{\theta}$) that the set of points $\hat{\boldsymbol{\theta}}(\boldsymbol{a})$ obtained in this manner describes a closed convex set in $\boldsymbol{\theta}$ which contains $\hat{\boldsymbol{\theta}}$ as an inner point. This closed convex set is our desired confidence region for $\boldsymbol{\theta}$ and it is readily seen that this also remains translation invariant.

In actual practice, the problem of finding out this convex close set in $\boldsymbol{\theta}$ (say $C(\boldsymbol{\theta})$), is not very complicated. It can be shown (cf. Sen [15]) that under the regularity conditions of section 2,

(3.21) $$n^{\frac{1}{2}}(h_j^*(\theta_j(\boldsymbol{a}))-\mu_N^{(j)}) \overset{\mathcal{L}}{\sim} \overline{B}_j(\boldsymbol{F})\,.n^{\frac{1}{2}}(\hat{\theta}_j(\boldsymbol{a})-\hat{\boldsymbol{\theta}}_j),\ j = 1, \ldots, p,$$

where $\overline{B}_j(F)$, $j = 1, \ldots, p$ are defined in (3.13). Thus, asymptotically $C(\boldsymbol{\theta})$ reduces to the following from

(3.22) $$C(\boldsymbol{\theta}) = \{\boldsymbol{\theta} : n(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\boldsymbol{\tau}^{-1}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})' \leqslant \chi^2_{p,\varepsilon}\},$$

where $\boldsymbol{\tau}$ is defined in (3.14). If we take any point $\boldsymbol{a}$ on the boundary of the ellipsoid $H_N(\boldsymbol{\theta}) \leqslant \chi^2_{p,\varepsilon}$, then using (3.19), (3.20) and (3.21), we have

$$(3.23) \qquad \bar{B}_j(F) \overset{P}{\sim} (a_j - \mu_N^{(j)})/(\hat{\theta}_j(\boldsymbol{a}) - \hat{\theta}_j), \quad j = 1, \ldots, p.$$

Thus, on considering one or more points $\boldsymbol{a}$ on the boundary of the ellipsoid $H_N(\boldsymbol{\theta}) \leqslant \chi^2_{p,\varepsilon}$, we can estimate $\bar{B}_j(\boldsymbol{F})$ and then using (3.14) and $\hat{\boldsymbol{\nu}}$ (defined just before (3.16),), we get $\hat{\boldsymbol{\tau}} = ((\hat{\tau}_{jl}))$, where $\hat{\tau}_{jl} = \hat{\nu}_{jl}/\bar{B}_j^*(F)\,\bar{B}_l^*(\boldsymbol{F})$, $j, l = 1, \ldots, p$, and $\bar{B}_j^*(\boldsymbol{F})$ denote the estimate obtained by (3.23). Once $\boldsymbol{\tau}$ is obtained, if we define a set $\hat{C}(\theta)$ by

$$(3.24) \qquad \hat{C}(\boldsymbol{\theta}) = \{\boldsymbol{\theta} : n(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})\,\hat{\boldsymbol{\tau}}^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})' \leqslant \chi^2_{p,\varepsilon}\},$$

then it is easily seen that $\hat{C}(\boldsymbol{\theta}) \overset{P}{\sim} C(\boldsymbol{\theta})$. Consequently, in actual practice, we may recommend (3.24) as the working confidence region for θ with confidence coefficient 1-ϵ. The simultaneous confidence region (3.22) or (3.24) remains valid for all $F_1, \ldots, F_r$ with covariance structures not necessarily identical. A word of clarification is necessary here. Though (3.24) is an asymptotic confidence region, in actual practice, if we select $\boldsymbol{h}^*$ some simple functions (e.g., rank-sum etc.), this remains valid for moderately large samples, where the usual parametric confidence-regions (under heterogeneity of dispersion matrices) may not be very simple.

It may be noted that if we borrow the idea of asymptotically smallest confidence region (cf. Wilks [18, pp 384-389]) then the asymptotic optimality or efficiency of the confidence region (3.22) or (3.24) with respect to the standard parametric confidence region for $\boldsymbol{\theta}$ (that may be derived with the help of likelihood ratio function) comes out to be

$$(3.25) \qquad \{|\bar{\boldsymbol{\Sigma}}|/|\boldsymbol{\tau}|\}^{\frac{1}{2}} = \{e(\boldsymbol{\tau}, \bar{\boldsymbol{\Sigma}})\}^{p/2},$$

where $\bar{\boldsymbol{\Sigma}}$, $\boldsymbol{\tau}$ and $e(\boldsymbol{\tau}, \bar{\boldsymbol{\Sigma}})$ are defined in (2.43), (3.14) and (3.15) respectively. Thus, the asymptotic optimality of the confidence region and the point estimate of $\boldsymbol{\theta}$ are functionally related to each other.

4. CHOICE OF RANK-ORDER TEST AND THE RELATED EFFICIENCY OF THE MANOVA PROCEDURES

We have so far considered a general class of rank-order statistics and developed some nonparametric MANOVA procedures. Now, we shall consider some specific rank-order statistics and study the related asymptotic efficiency factors. In particular, we shall consider the following three types of statistics, where,

$$\boldsymbol{E}_N = (E_{N1}, \ldots, E_{NN}), \quad E_{N\alpha} = J_N(\alpha/(N+1))\ \alpha = 1, \ldots, N;$$

are characterized as follows.

(1) *Median procedure*: Here

$$E_{N\alpha} = 1, \text{ if } \alpha \leqslant [N/2], \qquad = 0, \text{ otherwise.} \tag{4.1}$$

(2) *Rank-sum procedure.* Here

$$E_{N\alpha} = \alpha \text{ for } 1 \leqslant \alpha \leqslant N. \tag{4.2}$$

(3) *ψ-score procedure.* Let ψ be any specified c.d.f., and let $E_{N\alpha}$ be the expected value of the αth order statistic in a sample of size N drawn from a population with the c.d.f. ψ. In particular, if ψ is taken to be a standardized normal c.d.f, the procedure will be termed the *normal score procedure.*

In the univariate analysis of variance problem, it is known (cf. [6]), that against normal alternatives, the median procedure has an efficiency only $2/\pi$, though the same may be quite high for some non-normal c.d.f's. The asymptotic efficiency of rank-sum procedure for normal alternatives is $3/\pi$, and it has a lower bound .864 for any c.d.f. (cf [6]), though it can be arbitrarily large for some typical c.d.f. The normal score procedure has always an asymptotic efficiency greater than or equal to one, for all continuous c.d.f. (c.f. [7]).

In the multivariate case, it can be shown easily that all these procedures satisfy the regularity conditions (C.i), $i = 2, 3, 4, 5$. Further, if the c.d.f.'s $F_1, \ldots, F_r$ are non-singular in the sense that the cluster of the points is not confined in any $p-1$ dimensional subsequence of the p-dimensional Euclidean space, then it can be shown (cf. [10]) that (C.6) also holds. Regarding the expressions (2.42), (3.15) and (3.25), it can be shown that for normal $F_1, \ldots, F_r$ these are all equal to unity if we adopt the normal score procedure. Thus, the use of normal scores preserves the distribution-free property of the MANOVA procedure and at the same time, makes them asymptotically full efficient against normal alternatives. The efficiency factors can be shown to be greater than one for various non-normal c.d.f.'s but it can not be shown that they are greater than or equal to one for all $\boldsymbol{F} = F_1, \ldots, F_r)$ of the continuous type. Bickel [1] while considering the efficiency of Hodges-Lehmann [9] estimate of shift in the p-variate single sample case, came across a similar situation. He, however, considered only the rank-sum and median procedures. In the particular case of bivariate normal c.d.f., he deduced a lower bound (about .87) for the minimum efficiency of rank-sum procedure, while the minimum efficiency of the median procedure may be arbitrarily low. However, for more than two variates and/or for non-normal c.d.f.'s it is very difficult to prescribe any lower bound for the efficiency factors (2.42), (3.15) or (3.25), as they depend explicitly on the associated matrix $\boldsymbol{\tau}$, $\boldsymbol{\nu}$ and $\overline{\boldsymbol{\Sigma}}$, and nothing can be said, in general, about the magnitude or bounds for the characteristic roots of $\boldsymbol{\nu}\overline{\boldsymbol{\Sigma}}^{-1}$ or $\boldsymbol{\tau}\overline{\boldsymbol{\Sigma}}^{-1}$. However, if the p-variates are symmetrically dependent and have jointly a p-variate normal c.d.f. with a common correlation ρ, then it can be shown that the normal score procedure is better than the rank-sum procedure, which in turn is better than the median procedure.

In actual practice, the rank-sum procedure results in great simplification of the actual computation, while the normal score procedure is anticipated to have a better performance characteristic for nearly normal c.d.f.'s, and the choice will depend on the practical convenience and the degree of precision aimed at.

References

[1] Bickel, P. J. (1964). "On some Alternative Estimates for Shift in the p-Variate One Sample Problem," *Ann. Math. Statist.*, **35**, 1079–1090.

[2] Chatterjee, S. K. and Sen, P. K. (1964). "On Some Nonparametric Tests for the Bivariate Two Sample Location Problem," *Calcutta Statist. Asso. Bull.*, **13**, 18–58.

[3] Chatterjee, S. K. and Sen, P. K. (1969). "Nonparametric Tests for the Multisample Multivariate Location Problem," in *Contributions to Statistics and Probability, Essays in Memory of Samarendra Nath Roy*, University of North Carolina Press).

[4] Chernoff, H. and Savage, I. R. (1958). "Asymptotic Normality and Efficiency of Certain Nonparametric Test Statistics," *Ann. Math. Statist.*, **20**, 972–994.

[5] Govindarajulu, Z., Lecam, L. and Raghavachari, M. (1966). "Generalization of Chenoff-Savage Theorems on Asymptotic Normality of Nonparametric Test Statistics," *Proc. 5th Berkeley Symp. Math. Statist. Prob. Univ. of Calif.* 608-638.

[6] Hodges, J. L. Jr. and Lehmann, E. L. (1956). "The Efficiency of Some Nonparametric Competitors of the t-Test," *Ann. Math. Statist.*, **27**, 324–335.

[7] Hodges, J. L. Jr. and Lehmann, E. L. (1960). "Comparison of the Normal Scores and Wilcoxon Tests," *Proc. 4th Berkeley Symp. Math. Stat. Prob. Univ. of Calif.*, **1**, 83–92.

[8] Hodges, J. L. Jr. and Lehmann, E. L. (1962). "Rank Methods for Combination of Independent Experiments in Analysis of Variance," *Ann. Math. Statist.*, **33**, 482–497.

[9] Hodges, J. L. Jr. and Lehmann, E. L. (1963). "Estimates of Location Based on Rank Tests," *Ann. Math. Statist.*, **34**, 598–611.

[10] Puri, M. L. and Sen, P. K. (1966). On a Class of Multivariate Multisample Rank-Order Tests," *Sankhyā, Ser. A.*, **28**, 353–376.

[11] Sen, P. K. (1963). "On the Estimation of Relative Potency in Dilution (-Direct) Assays by Distribution-Free Methods," *Biometrics*, **19**, 532–552.

[12] Sen, P. K. (1965). "Some Further Applications of Non-parametric Methods in Dilution (-Direct) Assays," *Biometrics*, **21**, 799–810.

[13] Sen, P. K. (1966). "On a Class of Two Sample Bivariate Nonparametric Tests," *Proc. 5th Berkeley Symp. Math. Stat. Prob. Univ. of Calif.* 639-656.

[14] Sen, P. K. (1965). "On a Class of Multisample Multivariate Nonparametric Tests," (unpublished.)

[15] Sen, P. K. (1965). "On a Distribution-Free Method of Estimating Asymptotic Efficiency of a Class of Nonparametric Tests," *Ann. Math. Statist.*, **37**, 1759–1770.

[16] Sen, P. K. (1965). "On Pooled Estimation and Testing of Heterogeneity of Shift Parameters by Distributon-Free Methods. *Calcutta Statist. Assoc. Bull.* **16**, 139-157.

[17] Sverdrup, E. (1952). "The Limit Distribution of a Continuous Function of Random Variables," *Skand. Actdsfkt.*, **35**, 1–10.

[18] Wilks, S. S. (1962). *Mathematical Statistics*, John Wiley and Sons, New York.

(*Received Jan. 1, 1966.*)